

commit_message is_approved
before_it_is_merged there_is_a_bug
looks_good_to_me
someone_else_must_approve
this_needs_some_tweaks
makes_sense thanks_for should_be
there_is_a_problem would_be_better_to
uncaught_exception



Four Eyes Are Better Than Two: On the Impact of Code Reviews on Software Quality

GitリポジトリとGerritリポジトリを関連付けた
分析を通して、コードレビューがソフトウェア
品質に与える影響を調べた。

著者

Faculty of Computer Science
FreeUniversity of Bozen-Bolzano



Gabriele Bavota



Barbara Russo

背景

- ▷ 近年のコードレビュー活用
- ▷ コードレビューの影響の分析
 - コードレビューリポジトリのみ
- ▷ ソフトウェア品質全体との関係は分析されていなかった.



コードレビューとバグの誘発,
さらにはソフトウェア品質の
関係について調査する

Research Questions

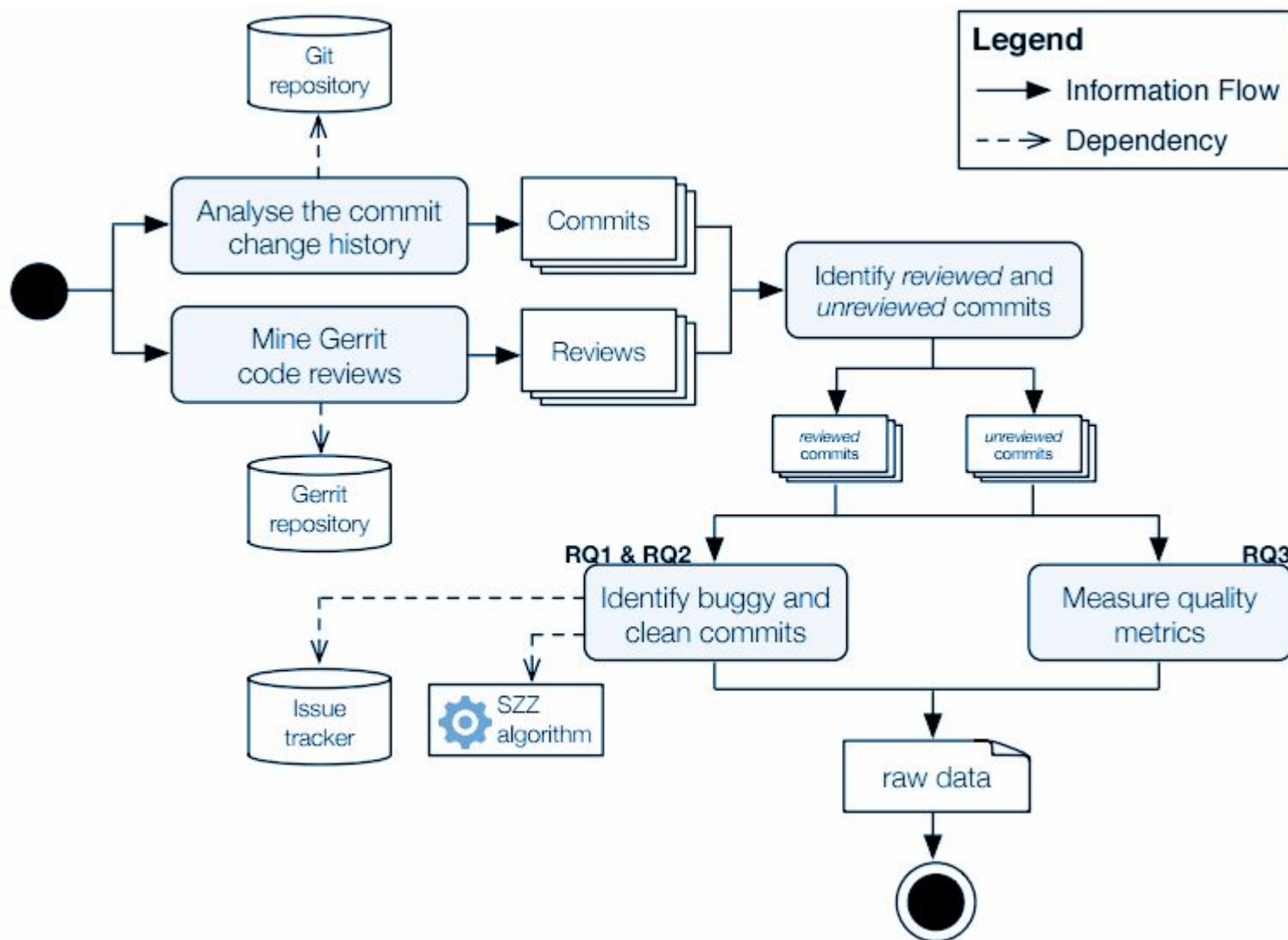
RQ1. コードレビューはバグの発生に影響を与えるか.

RQ2. コードレビューの参加者の数はバグの発生に影響を与えるか.

○ 参加者 = レビューした人 + コメントを残した人の数

RQ3. コードレビューはソフトウェア品質に影響を与えるか(コンポーネントレベル).

手法



対象プロジェクト



*Android frameworks
base APIs*



LibreOffice
The Document Foundation



CHARACTERISTICS OF SYSTEMS UNDER ANALYSIS

Project	Link	Language	Period	KLOC	#Issues	#Commits	
						Reviewed	Unreviewed
Android frameworks base APIs	http://tinyurl.com/lwoypdk	Java	Jan 2010-Dec 2014	534-1,043	9,311	1,593	31,761
LibreOffice	http://www.libreoffice.org/	C++	Mar 2011-Oct 2014	64-6,812	3,104	5,337	45,763
Scilab	http://www.scilab.org	Java/C++	Jan 2010-Sep 2014	679-3,172	1,421	2,499	243
Overall	-	-	-	-	13,836	9,429	77,677

※選定条件

- ▷ レビューアクティビティが追える
- ▷ 十分な数のレビューとアンレビューなコミットが存在する

RQ1 : 結果

System	NB_{NR}	B_{NR}	NB_R	B_R	p -value	OR
Android APIs	23,439	8,322	1,521	81	<0.001	3.98
LibreOffice	43,216	2,547	5,190	147	<0.001	2.08
Scilab	119	15	2,370	129	0.001	2.20

- ▷ 明らかにレビューしたほうがClean
 - マン・ホイットニーのU検定
 - フィッシャーの正確確率検定

◎レビューしていないコードは、
バグを生み出しやすい。

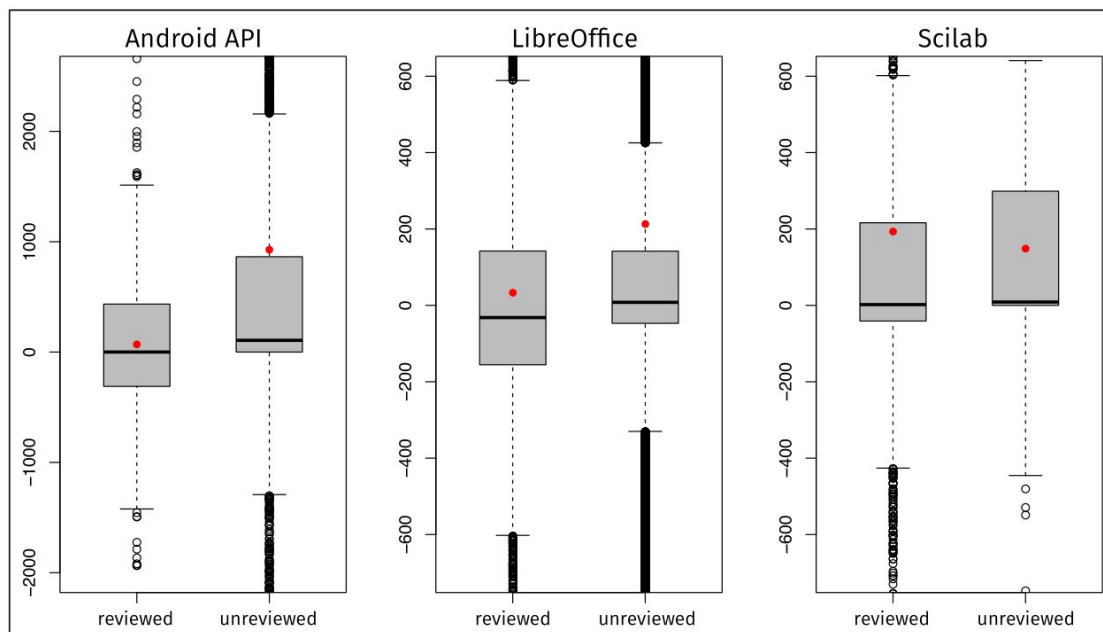
RQ2: 結果

System	Number of reviewers									
	Clean reviewed commits					Buggy reviewed commits				
	mean	median	min	max	st. dev.	mean	median	min	max	st. dev.
Android APIs	4.15	4.00	1.00	15.00	2.19	4.28	4.00	1.00	12.00	2.39
LibreOffice	1.28	1.00	1.00	6.00	0.60	1.08	1.00	1.00	3.00	0.31
Scilab	1.22	1.00	1.00	5.00	0.50	1.06	1.00	1.00	4.00	0.33

- ▷ クリーンなコミットはレビュアー数も多い
- 2/3のプロジェクトで支持

△コードレビューの参加者数が多い方が
バグが発生しにくい

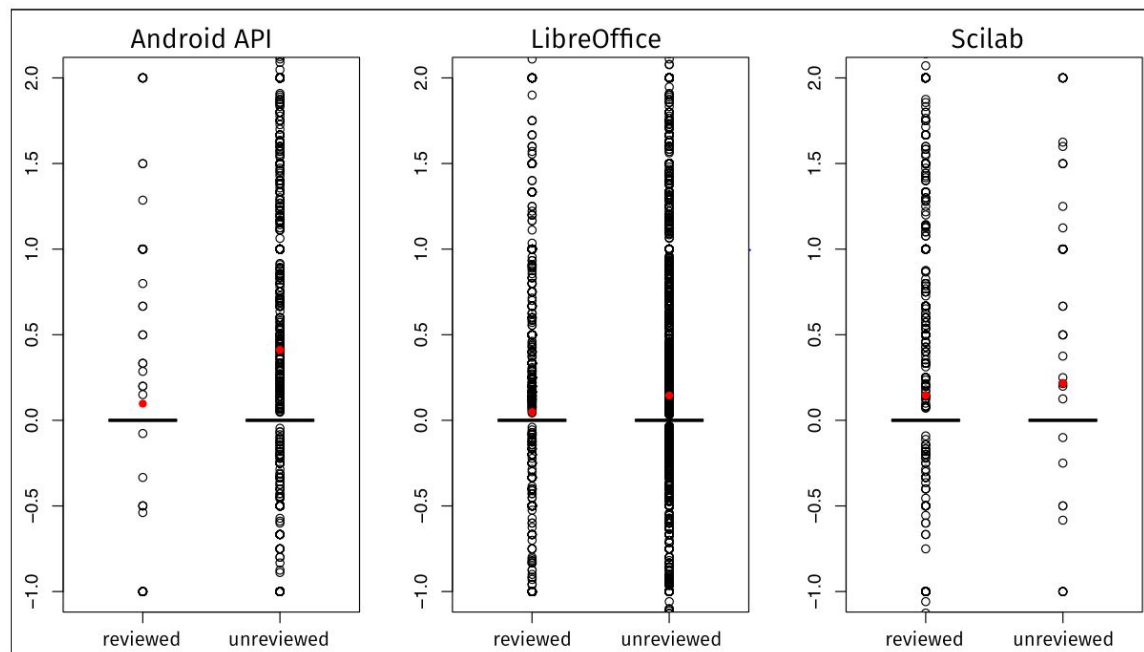
RQ3: 結果



レビューによるコード複雑度の推移

- ▷ 複雑度を少し小さくする.
- クリフの d

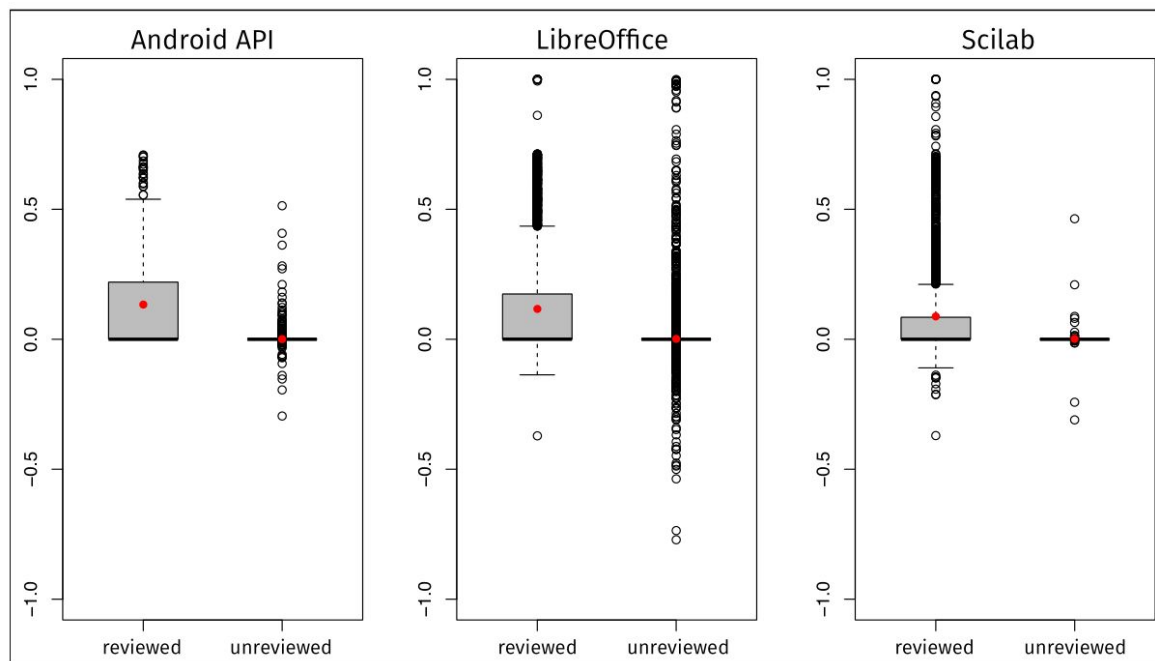
RQ3: 結果



レビューによるコードカップリングの推移

- ▷ コードカップリングに影響は与えない

RQ3: 結果



レビューによるコード可読性の推移

- ▷ 可読性は良くなる
- コーエンのd(標準偏差の変倍)

RQ3:本論の結果

- ▷ 複雑度(Complexity)を小さくする
- ▷ コードカップリングに影響なし
- ▷ 可読性(Readability)は良くなる

◎コードレビューはソフトウェアの品質を良くする.

本論のまとめ

- ▷ コードレビューとソフトウェア品質
 - レビューしないと, バグを生み出しやすい.
 - レビューの参加者は多いほうが良い
 - レビューすると, 品質を高く維持できる.
 - Clean, Readable, Brief

- ▷ 所感
 - 読みやすい論文だった(3年生の1本目に最適?)
 - 各結果に対してきちんと検定を行っていた
 -

Thanks!

Any questions?

You can find me at:

@username

user@mail.me

RQ1: おまけ

see_inline_comments
commit_message **is_approved**
before_it_is_merged there_is_a_bug
looks_good_to_me
someone_else_must_approve
this_needs_some_tweaks
makes_sense **thanks_for** should_be
there_is_a_problem would_be_better_to
uncaught_exception